

Infoblatt zur Sicherung, Archivierung, Veröffentlichung und Löschung von Forschungsdaten

2. August 2017

Öffentlich verfügbare und qualitativ gehaltvolle Forschungsdaten gehören zur „guten wissenschaftlichen Praxis“.¹ Sie erzeugen eine Transparenz, durch die sich wissenschaftliche Arbeiten und ihre Qualität nicht nur besser nachvollziehen lassen. Die dauerhafte Sicherung und Bereitstellung von Forschungsdaten ermöglichen es der Forschungsgemeinschaft darüber hinaus, auf dem Forschungsstand aufzubauen² sowie die Forschungsdaten für andere Forschungszwecke zu nutzen. Damit gelten qualitätsgesicherte Daten als „Grundpfeiler wissenschaftlicher Erkenntnis“.³ Verschiedene Institutionen, wie z. B. der Wissenschaftsrat, die Allianz der deutschen Wissenschaftsorganisationen oder die Deutsche Forschungsgemeinschaft (DFG), sprechen sich schon seit Jahren für eine dauerhafte und freie Verfügbarkeit von Forschungsdaten aus. Aus diesen Forderungen leiten sich für die ForscherInnen ggf. vertragliche Verpflichtungen im Umgang mit Forschungsdaten ab. Forschungsdaten werden im Verlauf von Forschungsprozessen generiert, z. B. durch Digitalisierung, Quellenforschung, Experimente, Messungen, Erhebungen oder Befragungen⁴. Sie können aus Messdaten, Laborwerten, audiovisuellen Informationen, Texten, Umfragedaten, Objekten aus Sammlungen oder Proben und vielem mehr bestehen. Auch die im Forschungsvorhaben verwendeten Methoden, wie z. B. Fragebögen, Software und Simulationen, können zu Forschungsdaten zählen, wenn sie zentrale Forschungsergebnisse darstellen.⁵

Neben den wissenschaftsethischen und vertraglichen Gründen zum verantwortungsvollen Umgang mit empirischen Daten sind auch gesetzliche Vorgaben zu beachten. Aus datenschutzrechtlichen Gründen bedürfen insbesondere personenbezogene Daten erhöhter Aufmerksamkeit.

Im Folgenden soll ohne Anspruch auf Vollständigkeit ein Überblick über die Standards bei der Sicherung, Archivierung und Veröffentlichung von Forschungsdaten gegeben werden, an den sich ein Blick auf die gesetzlichen Grundlagen des Datenschutz- und Urheberrechts

¹http://www.allianzinitiative.de/fileadmin/user_upload/www.allianzinitiative.de/Grundsaeetze_Forschungsdaten_2010.pdf.

²http://www.dfg.de/foerderung/antragstellung_begutachtung_entscheidung/antragstellende/antragstellung/nachnutzung_forschungsdaten/; Zugriff am 9. Februar 2017.

³http://www.allianzinitiative.de/fileadmin/user_upload/www.allianzinitiative.de/Grundsaeetze_Forschungsdaten_2010.pdf; Zugriff am 9. Februar 2017.

⁴ <http://www.allianzinitiative.de/de/handlungsfelder/forschungsdaten/>; Zugriff am 9. Februar 2017.

⁵ <http://www.allianzinitiative.de/de/handlungsfelder/forschungsdaten/>; Zugriff am 9. Februar 2017.

anschließt. Dabei wird die besondere Rolle und Handhabung von personenbezogenen Daten berücksichtigt. Da die Konventionen im Umgang mit Forschungsdaten nach Disziplinen, Institutionen und Einheiten stark variieren,⁶ kann die nachfolgende Darstellung nur einem ersten Überblick dienen. Insbesondere ersetzt sie nicht die gründliche Einholung von Spezialinformationen zum eigenen Forschungsbereich.

1. Sichern

Forschungsdaten müssen zur Einhaltung **wissenschaftlicher** Maßstäbe oftmals bis zu 10 Jahren gespeichert werden.⁷ Soweit für die Forschung im Forschungszeitraum notwendig (siehe 3.), können auch personenbezogene Daten auf den eigenen Geräten gespeichert werden. Dabei sollten die Daten optimaler Weise auf mehreren Wegen gesichert werden, wie z. B. zusätzlich auf einer externen Festplatte. Es muss hierbei jedoch beachtet werden, dass unbefugten Dritten kein Zugriff gewährt wird. Aus diesem Grund ist eine Speicherung personenbezogener Daten in der Cloud nach Möglichkeit zu vermeiden. Es ist zwar auch möglich, die Sicherung der Daten an Dienstleister zu übergeben. Sobald die Daten jedoch aus der Hand gegeben werden, müssen personenbezogene Daten anonymisiert werden.

Von vielen Fördermittelgebern werden konkrete **Speicherungszeiträume** vorgegeben, an die die ForscherInnen gebunden sind.⁸ Insofern ist darauf hinzuweisen, dass zu den Forschungsdaten nicht etwa nur die reinen, in der Regel als Zahlen vorliegenden Ergebnisse gehören, sondern auch weitere Daten, die der Erhebung und Auswertung dienen bzw. dienen. Dazu gehören (ausgefüllte) Fragebögen, Interviewleitfäden, sonstige Rückmeldungen und Notizen und die verwendete Software bzw. das Programm sowie der Code, die zur Analyse herangezogen wurden. Alle diese Daten sind wichtig, um die Forschung im Falle von Kritik an der Datenerhebung und -auswertung adäquat verteidigen zu können. Deshalb – sowie für die Weiterverwendbarkeit der Daten – sollten nach Möglichkeit auch die Roh- und Metadaten über die von der DFG vorgeschlagenen zehn Jahre hinaus gespeichert werden.

2. Archivieren und Veröffentlichen

Zur Veröffentlichung von Forschungsdaten gehört neben der eigentlichen Publikation auch eine Beschreibung der Daten. In der Regel werden die veröffentlichten Daten mit einem

⁶ Für einen Überblick siehe: http://www.forschungsdaten.org/index.php/Data_Policies; Zugriff am 9. Februar 2017.

⁷ http://www.dfg.de/download/pdf/foerderung/programme/lis/ua_inf_empfehlungen_200901.pdf.

⁸ vgl. <https://de.slideshare.net/heinzpampel/vorlesung-forschungsdatenrepositorien-ss14>.

Persistent Identifier (wie einer DOI oder eines Handles) versehen und dadurch eindeutig referenzier- und zitierbar.⁹

Schon 2001 hob der Wissenschaftsrat hervor, wie wichtig es ist, dass digitale Informationen für die Wissenschaft verfügbar bleiben, und forderte die Wissenschafts- und Förderorganisationen dazu auf, ihre Daten zu archivieren und für die weitere Nutzung in der Wissenschaft verfügbar zu halten. Der Wissenschaftsrat betonte dabei weiterhin die teilweise divergierenden internen Standards unterschiedlicher Fächerkulturen sowie medienspezifische Besonderheiten.

Die Allianz der deutschen Wissenschaftsorganisationen veröffentlichte im Jahr 2010 Grundsätze zum Umgang mit Forschungsdaten. Als zentrale fächerübergreifende Gemeinsamkeiten geben die Grundsätze vor allem die **langfristige Sicherung** und **öffentliche Zugänglichkeit** der Daten, die Bedeutung eines ausreichenden Weiterbildungsangebots zum professionellen Datenmanagement, die Verwendung gemeinsamer Standards in der Dokumentation im Sinne einer transdisziplinären Nutzung und die Entwicklung weitreichender technischer und organisatorischer Infrastruktur an.

Zu beachten ist schließlich, dass bei den verschiedenen Institutionen in der Regel nur von einer Veröffentlichung für die Forschung die Rede ist, aber gegebenenfalls auch eine generell öffentliche Publikation als Veröffentlichungsformat in Frage kommt.

3. Gesetzliche Grundlagen

a) Datenschutzrecht

Wie eingangs erwähnt fallen Forschungsdaten, die einen Bezug zu einer bestimmten oder bestimmbarer Person aufweisen, in den Bereich des Datenschutzrechts. Als **bestimmbar** wird eine Person angesehen, die direkt oder indirekt, insbesondere mittels Zuordnung zu einer Kennung wie einem Namen, einer Kennnummer, Standortdaten, einer Online-Kennung oder zu einem oder mehreren besonderen Merkmalen, die Ausdruck der Identität dieser natürlichen Person sind, identifiziert werden kann.¹⁰ Das Datenschutzrecht hat das Ziel, **sämtliche Informationen** zu erfassen, die mit einer Person in Verbindung gebracht werden können. Sind personenbezogene Daten betroffen, unterfallen sämtliche Erhebungs-, Verarbeitungs- und Nutzungsvorgänge dem Datenschutzrecht.

Neben dem primär einschlägigen Bundesdatenschutzgesetz (BDSG) existiert in jedem Bundesland ein Landesdatenschutzgesetz, das insbesondere die Datenverarbeitung durch öffentliche Stellen des Landes regelt. Damit unterfällt die Forschung an **öffentlich-rechtlichen Universitäten** in aller Regel dem jeweiligen **Landesdatenschutzrecht**, welches

⁹ <http://www.forschungsdaten.org/index.php/Datenpublikation>.

¹⁰ Hier wie im Folgenden: BeckOK/Schild, § 1 BDSG Rn. 1a.

sich allerdings oftmals am BDSG orientiert. Zudem findet das BDSG auf die Forschung an **privaten Forschungseinrichtungen** und Einrichtungen des **Bundes** unmittelbare Anwendung. Nach Bundesrecht wie Landesrecht ist die Erhebung, Nutzung und Verarbeitung personenbezogener Daten grundsätzlich unzulässig, es sei denn, die betroffene Person willigt in die Nutzung ihrer Daten ein oder es besteht eine gesetzliche Erlaubnis (§ 4 Abs. 1 BDSG). Eine **Einwilligung** des Betroffenen muss **freiwillig, informiert** und im Regelfall **schriftlich** abgegeben werden. Handelt es sich um sensible persönliche Daten des Betroffenen (z. B. mit Bezug zur Gesundheit oder zur politischen/religiösen Ausrichtung), muss die Einwilligung zudem **ausdrücklich**, d. h. gerade in Bezug auf die Verarbeitung dieser Daten, erfolgen.

Selbst bei Vorliegen einer wirksamen Einwilligung zur Verarbeitung nach den genannten Voraussetzungen sind personenbezogene Forschungsdaten zu anonymisieren, sobald dies nach dem Forschungszweck möglich ist (§ 40 Abs. 2 S. 1 BDSG). Bei einer solchen **Anonymisierung** ist der Bezug zur konkreten Person derart zu entfernen, dass die Daten dem Betroffenen endgültig nicht mehr zugeordnet werden können. Eine Anonymisierung wird nicht erst dann erforderlich, wenn das Forschungsprojekt abgeschlossen ist, sondern schon dann, wenn die Daten nicht mehr für die Forschung benötigt werden.¹¹ Solange der Personenbezug für die Forschung unerlässlich ist, sind die Daten zumindest zu **separieren**, d. h. die Merkmale, die zur Bestimmung der Person führen können, sind getrennt von den erhobenen Forschungsdaten zu verwahren. Dabei darf eine erneute Zusammenführung nur erfolgen, wenn, soweit und solange der Forschungszweck dies erforderlich macht.

Eine **Veröffentlichung** der **Forschungsergebnisse** ist uneingeschränkt zulässig, wenn diese keine personenbezogenen Daten (mehr) enthalten und sich auch kein Bezug zu solchen mehr herstellen lässt. Dies gilt selbst dann, wenn die Forschungsergebnisse datenschutzrechtswidrig zustande gekommen sind.¹² Lässt die Veröffentlichung noch einen Rückschluss auf personenbezogene Daten zu oder werden diese sogar mitveröffentlicht, bedarf es dagegen einer (gesonderten, im Zweifel auch erneuten) **Einwilligung** des Betroffenen (§ 40 Abs. 3 BDSG). Eine eng begrenzte Ausnahme gilt für den Bereich der historischen Forschung, sofern die Forschung Ereignisse der Zeitgeschichte betrifft und die Veröffentlichung der personenbezogenen Daten zur Darstellung der Ergebnisse unerlässlich ist.

¹¹ BeckOK/Lindner, § 40 BDSG Rn. 25 ff.

¹² BeckOK/Lindner, § 40 BDSG Rn. 29.

b) Urheberrecht

Neben dem Datenschutzrecht vermag auch das Urheberrecht die Materie des Forschungsdatenmanagements zu beeinflussen. Im Unterschied zum Datenschutzrecht greift der Urheberrechtsschutz (§ 11 UrhG) allerdings erst, sobald aus den Daten ein „**Werk**“ geschaffen wird. Dies ist der Fall, wenn aus ihnen eine persönliche geistige Schöpfung gebildet wird (§ 2 Abs. 2 UrhG).

Ein solches Werk kann durch die (besondere) Zusammenstellung, Darstellung oder Auswertung von Forschungsdaten geschaffen werden. Liegt ein geschütztes Werk vor, so kann es ohne Zustimmung des Urheberrechtinhabers nur zum rein privaten Gebrauch vervielfältigt oder zur **Zitierung** bzw. Vervielfältigung bloßer **Tatsachen** im Rahmen einer eigenen Interpretation verwendet werden. Gleiches gilt für **Datenbanken**, an denen der Ersteller der Datenbank das alleinige Verbreitungs- und Vervielfältigungsrecht hat.¹³

Die aus dem Urheberrecht erwachsenden Schutzrechte sind zunächst bei der Nutzung **fremder und eigener** Forschungsdaten zu berücksichtigen. Eine Verwendung fremder Werke und Datenbanken zum Gebrauch unter Ausschluss der Weitergabe an Dritte ist zulässig, wenn sie zum privaten wissenschaftlichen Zweck geboten ist. Hiervon erfasst ist das Erstellen von **Kopien**, wenn der Kauf eines Originals nicht (wirtschaftlich) zumutbar oder mit faktischen Schwierigkeiten verbunden ist.¹⁴ Das Urheberrecht an Werken erlischt i. d. R. 70 Jahre nach dem Tod des Schöpfers.

In seltenen, primär technischen Forschungsbereichen kann auch das **Patentrecht** Rechte an Forschungsdaten vermitteln oder beschränken. Zur Vermeidung von Streitfällen sollte im Zweifel stets eine Klärung mit den Rechteinhabern bzw. den Zugriffsinteressierten angestrebt werden, die in einen Lizenzierungs- oder Überlassungsvertrag münden kann.

¹³ Leitfaden zum Forschungsdatenmanagement, Handreichungen aus dem WissGrid-Projekt, 2013, S. 57.

¹⁴ Leitfaden zum Forschungsdatenmanagement, Handreichungen aus dem WissGrid-Projekt, 2013, S. 58.